

# Non-backtracking random walk

(Preprint)

Robert Fitzner\*

Remco van der Hofstad\*

December 11, 2012

## Abstract

We consider non-backtracking random walk (NBW) in the nearest-neighbor setting on the  $\mathbb{Z}^d$ -lattice and on tori. We evaluate the eigensystem of the  $m \times m$ -dimensional transition matrix of NBW where  $m$  denote the degree of the graph. We use its eigensystem to show a functional central limit theorem for NBW on  $\mathbb{Z}^d$  and to obtain estimates on the convergence towards the stationary distribution for NBW on the torus.

## 1 Introduction

The non-backtracking walk (NBW) is a simple random walk that is conditioned not to jump back along the edge it has just traversed. NBW can be viewed as a Markov chain on the set of directed edges, where the edge has the interpretation of being the last edge traversed by the NBW, but we will not rely on this interpretation. We study NBWs on  $\mathbb{Z}^d$  in the nearest-neighbor setting, and NBWs on tori in various settings, derive the transition matrix for the NBW and analyse its eigensystem in Fourier space. We use this to study asymptotic properties of the NBW, such as its Green's function and a functional central limit theorem (CLT) on  $\mathbb{Z}^d$ , and its convergence towards the stationary distribution on the torus. In particular, our analysis allows us to give an explicit formula for the Fourier transform of the number of  $n$ -step NBWs traversing fixed positions at given times. We use this to prove that the finite-dimensional distributions of the NBW displacements, after diffusive rescaling, converge to those of Brownian motion. By an appropriate tightness argument, this proves a functional CLT. We further evaluate the Fourier transform of the NBW  $n$ -step transition probabilities on the torus to identify when the NBW transition probabilities are close to the stationary distribution. Our paper is inspired by the study of various high-dimensional statistical mechanical models. For example, our derivation allows us to give detailed estimates of the probability that NBW on a torus is at a given vertex, a fact used in the analysis of hypercube percolation in [6].

NBW have been investigated on finite graphs in [3], in particular expanders, where it was shown that NBWs mix faster than ordinary random walks. See also [4], where a Poisson limit was proved for the number of visits of an  $n$ -step NBW to a vertex in an  $r$ -regular graph of size  $n$ . In the nearest-neighbor setting on  $\mathbb{Z}^d$ , NBWs were investigated in [9, Section 5.3], where an explicit expression of its Green's function and many related properties are derived (see also [13, Exercise 3.8]). Noonan [10] investigates the generating function of NBWs, and his results also apply to walks that avoid their last 7 previous positions (i.e., with memory up to 8), and were used to improve the known upper bounds on the SAW connective constant. See [12] for an extension up to memory 22 for  $d = 2$ , further improving the upper bound on the SAW connective constant. The methods in [10, 12] allow to compute the generating function of the number of memory- $k$  SAWs for appropriate values of  $k$ , but do not investigate the number

---

\*Department of Mathematics and Computer Science, Eindhoven University of Technology, 5600 MB Eindhoven, The Netherlands. R.J.Fitzner@tue.nl, rhofstad@win.tue.nl

of memory- $k$  SAWs ending in a particular position in  $\mathbb{Z}^d$ . Finally, [11] studies relations between the exponential growth of the transition probabilities of NBW and non-amenability of the underlying graph. For interesting connections to zeta functions on graphs, which allow one to study the number of NBWs of arbitrary length ending in the starting point, we refer the reader to [7].

This paper is organized as follows. In Section 2 we investigate NBW on  $\mathbb{Z}^d$  and in Section 3 we study NBW on tori.

## 2 Non-backtracking random walk on $\mathbb{Z}^d$

### 2.1 The setting

An  $n$ -step nearest-neighbor simple random walk (SRW) on  $\mathbb{Z}^d$  is an ordered  $(n+1)$ -tuple  $\omega = (\omega_0, \omega_1, \omega_2, \dots, \omega_n)$ , with  $\omega_i \in \mathbb{Z}$  and  $\|\omega_i - \omega_{i+1}\|_1 = 1$ , where  $\|x\|_1 = \sum_{i=1}^d |x_i|$ . We always take  $\omega_0 = (0, 0, \dots, 0)$ . The step distribution of SRW is given by

$$D(x) = \frac{1}{2d} \mathbb{1}_{\{\|x\|_1=1\}}, \quad (2.1)$$

where  $\mathbb{1}_A$  is the indicator of the event  $A$ . If an  $n$ -step SRW  $\omega$  additionally satisfies that  $\omega_i \neq \omega_{i+2}$ , then we call the walk a *non-backtracking walk* (NBW). As the problem of NBW is trivial for  $d = 1$ , we always assume that  $d \geq 2$ . For the NBW we also count walks conditioned not to take their first step in a certain direction  $\iota$ . We exclusively use the Greek letters  $\iota$  and  $\kappa$  for values in  $\{-d, -d+1, \dots, -1, 1, 2, \dots, d\}$  and denote by  $e_\iota \in \mathbb{Z}^d$  the unit vector in direction  $\iota$ , i.e.  $(e_\iota)_\kappa = \text{sign}(\iota) \delta_{|\iota|, \kappa}$  (beware of the minus sign when  $\iota$  is negative, which is somewhat different from the usual choice of a unit vector). Let  $b_n(x)$  be the number of  $n$ -step NBWs with  $\omega_n = x$ , and  $b_n^\iota(x)$  the number of  $n$ -step NBWs  $\omega$  with  $\omega_n = x$  and  $\omega_1 \neq e_\iota$ . For  $n \geq 1$ , the following relations between these objects hold:

$$b_n(x) = \sum_{\iota \in \{\pm 1, \dots, \pm d\}} b_{n-1}^{-\iota}(x - e_\iota), \quad (2.2)$$

$$b_n(x) = b_n^\iota(x) + b_{n-1}^{-\iota}(x - e_\iota) \quad \forall \iota, \quad (2.3)$$

$$b_n^\iota(x) = \sum_{\kappa \in \{\pm 1, \dots, \pm d\} \setminus \{\iota\}} b_{n-1}^{-\kappa}(x - e_\kappa). \quad (2.4)$$

We analyse  $b_n$  and  $b_n^\iota$  using *Fourier theory*. For an absolutely summable function  $f: \mathbb{Z}^d \mapsto \mathbb{C}$ , we define its Fourier transform by

$$\hat{f}(k) = \sum_{x \in \mathbb{Z}^d} f(x) e^{ik \cdot x} \quad (k \in [-\pi, \pi]^d), \quad (2.5)$$

where  $k \cdot x = \sum_{s=1}^d k_s x_s$  and  $i$  denotes the imaginary unit. We use  $k$  exclusively to denote values in the Fourier dual space  $[-\pi, \pi]^d$ . For  $f, g: \mathbb{Z}^d \mapsto \mathbb{C}$  we denote their convolution by  $f \star g$ , i.e.,

$$(f \star g)(x) = \sum_{y \in \mathbb{Z}^d} f(y) g(x - y), \quad (2.6)$$

and note that the Fourier transform of  $f \star g$  is given by  $\hat{f}\hat{g}$ . Applying the Fourier transformation to (2.1)-(2.4) yields

$$\hat{D}(k) = \frac{1}{d} \sum_{i=1}^d \cos(k_i), \quad (2.7)$$

$$\hat{b}_n(k) = \sum_{\iota \in \{\pm 1, \dots, \pm d\}} \hat{b}_{n-1}^{-\iota}(k) e^{ik_\iota}, \quad (2.8)$$

$$\hat{b}_n(k) = \hat{b}_n^\iota(k) + \hat{b}_{n-1}^{-\iota}(k) e^{ik_\iota} \quad \forall \iota, \quad (2.9)$$

$$\hat{b}_n^\iota(k) = \sum_{\kappa \in \{\pm 1, \dots, \pm d\} \setminus \{\iota\}} \hat{b}_{n-1}^{-\kappa}(x) e^{ik_\kappa}. \quad (2.10)$$

In our further analysis, we use  $\mathbb{C}^{2d}$ -valued and  $\mathbb{C}^{2d} \times \mathbb{C}^{2d}$ -valued functions. For a clear distinction between scalar-, vector- and matrix-valued quantities, we always write  $\mathbb{C}^{2d}$ -valued functions with a vector arrow (e.g.  $\vec{v}$ ) and matrix-valued functions with bold capital letters (e.g.  $\mathbf{M}$ ). We do not use  $\{1, 2, \dots, 2d\}$  for the indices of the elements of a vector or a matrix, but use  $\{-d, -d+1, \dots, -1, 1, 2, \dots, d\}$  instead. Here, for a negative index  $\iota \in \{-d, -d+1, \dots, -1\}$  and a vector  $k \in [-\pi, \pi]^d$ , we define  $k_\iota := -k_{|\iota|}$ .

We denote the identity matrix by  $\mathbf{I} \in \mathbb{C}^{2d \times 2d}$  and the all-one vector by  $\vec{1} = (1, 1, \dots, 1)^T \in \mathbb{C}^{2d}$ . Moreover we define the matrices  $\mathbf{C}, \mathbf{J} \in \mathbb{C}^{2d \times 2d}$  by  $(\mathbf{C})_{\iota, \kappa} = 1$  and  $(\mathbf{J})_{\iota, \kappa} = \delta_{\iota, -\kappa}$ . To characterize the displacement of a step in a direction  $\iota$ , we define the diagonal matrix  $\hat{\mathbf{D}}(k)$  with the entries  $(\hat{\mathbf{D}}(k))_{\iota, \iota} = e^{ik_\iota}$ . We define the vector  $\vec{b}_n(k)$  with entries  $(\vec{b}_n(k))_\iota = \hat{b}_n^\iota(k)$ . Then, we can rewrite (2.8)-(2.10) as

$$\hat{b}_n(k) = \vec{1}^T \hat{\mathbf{D}}(-k) \vec{b}_{n-1}(k), \quad (2.11)$$

$$\hat{b}_n(k) \vec{1} = \vec{b}_n(k) + \hat{\mathbf{D}}(k) \mathbf{J} \vec{b}_{n-1}(k), \quad (2.12)$$

$$\vec{b}_n(k) = (\mathbf{C} - \mathbf{J}) \hat{\mathbf{D}}(-k) \vec{b}_{n-1}(k) = \left( (\mathbf{C} - \mathbf{J}) \hat{\mathbf{D}}(-k) \right)^n \vec{1}. \quad (2.13)$$

We define the transition matrix

$$\hat{\mathbf{A}}(k) = (\mathbf{C} - \mathbf{J}) \hat{\mathbf{D}}(-k), \quad (2.14)$$

so that  $(\hat{\mathbf{A}}(k))_{\iota, \kappa} = e^{-ik_\iota} (1 - \delta_{\iota, -\kappa})$ . With this notation in hand, we are ready to identify the NBW Green's function.

## 2.2 The Green's function

We start by deriving a formula for the NBW Green's function using the relations in (2.11)-(2.11). While these results are not novel, the analysis presented here is efficient and simple. We define the NBW Green's function as the generating function of  $\hat{b}_n$  and  $\hat{b}_n^\iota$ :

$$\hat{B}_z(x) := \sum_{n=0}^{\infty} \hat{b}_n(k) z^n, \quad \hat{B}_z^\iota(x) := \sum_{n=0}^{\infty} \hat{b}_n^\iota(k) z^n \quad (2.15)$$

$$\vec{B}_z(k)^T := (\hat{B}_z^1(k), \hat{B}_z^{-1}(k), \hat{B}_z^2(k), \dots, \hat{B}_z^{-d}(k)), \quad (2.16)$$

where  $\vec{y}^T$  denotes the transpose of the vector  $\vec{y} \in \mathbb{R}^d$ . By (2.11)-(2.12),

$$\hat{B}_z(k) = 1 + z \vec{1}^T \hat{\mathbf{D}}(-k) \vec{B}_z(k), \quad (2.17)$$

$$\hat{B}_z(k) \vec{1} = \vec{B}_z(k) + z \hat{\mathbf{D}}(k) \mathbf{J} \vec{B}_z(k) \Rightarrow \vec{B}_z(k) = \left[ \mathbf{I} + z \hat{\mathbf{D}}(k) \mathbf{J} \right]^{-1} \vec{1} \hat{B}_z(k), \quad (2.18)$$

Using  $\hat{\mathbf{D}}(k) \mathbf{J} \hat{\mathbf{D}}(k) \mathbf{J} = \mathbf{I}$ , it is easy to check that

$$\left[ \mathbf{I} + z \hat{\mathbf{D}}(k) \mathbf{J} \right]^{-1} = \frac{1}{1 - z^2} \left( \mathbf{I} - z \hat{\mathbf{D}}(k) \mathbf{J} \right), \quad (2.19)$$

and we use (2.17)-(2.19) to obtain

$$\hat{B}_z(k) = \frac{1}{1 - z\vec{1}^T \hat{\mathbf{D}}(-k) \left[ \mathbf{I} + z\hat{\mathbf{D}}(k)\mathbf{J} \right]^{-1} \vec{1}} = \frac{1 - z^2}{1 + (2d - 1)z^2 - 2dz\hat{D}(k)}. \quad (2.20)$$

Note in particular that

$$\hat{B}_z(k) = \frac{1 - z^2}{1 + (2d - 1)z^2} \hat{C}_{\mu_z}(k), \quad (2.21)$$

where  $\mu_z = 2dz/(1 + (2d - 1)z^2)$  and  $\hat{C}_\mu(k) = 1/[1 - \mu\hat{D}(k)]$  is the SRW Green's function. By (2.18),

$$\vec{B}_z(k) = \frac{1}{1 - z^2} \left[ \mathbf{I} - z\hat{\mathbf{D}}(k)\mathbf{J} \right] \vec{1} \hat{B}_z(k) = \frac{\left[ \mathbf{I} - z\hat{\mathbf{D}}(k) \right] \vec{1}}{1 + (2d - 1)z^2 - 2dz\hat{D}(k)},$$

so that

$$\hat{B}_z^\iota(k) = \frac{1 - ze^{ik_\iota}}{1 + (2d - 1)z^2 - 2dz\hat{D}(k)}. \quad (2.22)$$

### 2.3 The transition matrix

**The eigensystem.** We start by evaluating the transition matrix (2.14) by characterizing its eigenvalues and eigenvectors:

**Lemma 2.1** (Dominant eigenvalues). *For  $d \geq 2$  and  $k \in (-\pi, \pi)^d$ , let  $\hat{\mathbf{A}}(k)$  be the matrix given in (2.14). Then*

$$\hat{\lambda}_{\pm 1} = \hat{\lambda}_{\pm 1}(k) = d\hat{D}(k) \pm \sqrt{(d\hat{D}(k))^2 - (2d - 1)} \quad (2.23)$$

*are eigenvalues of  $\hat{\mathbf{A}}(k)$ . For  $k \neq (0, 0, \dots, 0)^T$ , the right eigenvectors  $\vec{v}^{(\pm 1)}$  to the eigenvalue  $\hat{\lambda}_{\pm 1}$  are given by*

$$\vec{v}^{(\pm 1)} = \hat{\lambda}_{\pm 1} \vec{1} \pm \hat{\mathbf{D}}(k) \vec{1}. \quad (2.24)$$

*For  $k = (0, 0, \dots, 0)^T$ , the eigenvectors are given by  $\vec{v}^{(1)}(0) = (2d - 2)\vec{1}$  and  $\vec{v}^{(-1)} := (1, -1, 0, 0, \dots, 0) \in \mathbb{Z}^{2d}$ .*

As we will see below, Lemma 2.1 yields the two most important eigenvalues. When  $\hat{\lambda}_+ = \hat{\lambda}_-$ , which occurs when  $(d\hat{D}(k))^2 - (2d - 1) = 0$ , it turns out that  $\vec{v}^{(1)}$  has geometric multiplicity 1, and that  $\vec{1}$  is a generalized eigenvector satisfying  $\hat{\mathbf{A}}(k)\vec{1} = \vec{v}^{(1)} + \hat{\lambda}_+\vec{1}$ . We continue by computing the remaining eigenvalues and -vectors:

**Lemma 2.2** (Simple eigenvalues). *For  $d \geq 2$ ,  $k \in (-\pi, \pi)^d$ , let  $\hat{\mathbf{A}}(k)$  be the matrix given in (2.14) and  $\vec{e}_\iota \in \mathbb{C}^{2d}$  the  $\iota$ th unit vector, i.e.,  $(\vec{e}_\iota)_\kappa = \delta_{\iota, \kappa}$  for  $\kappa \in \{\pm 1, \dots, \pm d\}$ . For  $\iota \in \{2, 3, \dots, d\}$ , let*

$$\begin{aligned} \vec{v}^{(\iota)} &= (1 + e^{ik_\iota})(e^{ik_\iota} \vec{e}_1 + \vec{e}_{-1}) - (1 + e^{ik_1})(e^{ik_\iota} \vec{e}_\iota + \vec{e}_{-\iota}) & \forall k \in [-\pi, \pi]^d \\ \vec{v}^{(-\iota)} &= (1 - e^{ik_\iota})(e^{ik_1} \vec{e}_1 - \vec{e}_{-1}) - (1 - e^{ik_1})(e^{ik_\iota} \vec{e}_\iota - \vec{e}_{-\iota}) & \forall k \in [-\pi, \pi]^d \text{ if } e^{ik_1} \neq 1, \end{aligned}$$

*Then,  $\vec{v}^{(\pm \iota)}$  is an eigenvector of  $\hat{\mathbf{A}}(k)$  to the eigenvalue  $\mp 1$ . Both eigenvalues have a geometrical multiplicity of  $d - 1$  for all  $k$ .*

Lemmas 2.1–2.2 identify a collection of  $2d$  independent eigenvectors, and thus the complete eigensystem, of  $\hat{\mathbf{A}}(k)$ . Now we prove these two lemmas:

*Proof of Lemma 2.1.* Let  $\hat{\lambda} \in \{\hat{\lambda}_1, \hat{\lambda}_{-1}\}$  and  $\vec{v} = \hat{\lambda}\mathbf{1} - \hat{\mathbf{D}}(k)\mathbf{1}$ . The values  $\hat{\lambda}_1$  and  $\hat{\lambda}_{-1}$  are the solutions of the quadratic equation

$$\hat{\lambda}^2 = 2d\hat{\lambda}\hat{D}(k) - (2d - 1). \quad (2.25)$$

Using  $\mathbf{C}\hat{\mathbf{D}}(-k)\vec{\mathbf{1}} = 2d\hat{D}(k)\vec{\mathbf{1}}$  and  $\mathbf{J}\hat{\mathbf{D}}(-k) = \hat{\mathbf{D}}(k)\mathbf{J}$ , we compute

$$\begin{aligned} \hat{\mathbf{A}}(k)\vec{v} &= (\mathbf{C} - \mathbf{J})\hat{\mathbf{D}}(-k)(\hat{\lambda}\mathbf{1} - \hat{\mathbf{D}}(k))\vec{\mathbf{1}} = \left(2d\hat{\lambda}\hat{D}(k)\mathbf{I} - \hat{\lambda}\hat{\mathbf{D}}(k) - (2d - 1)\mathbf{I}\right)\vec{\mathbf{1}} \\ &\stackrel{(2.25)}{=} \left(\hat{\lambda}^2\mathbf{I} - \hat{\lambda}\hat{\mathbf{D}}(k)\right)\vec{\mathbf{1}} = \hat{\lambda}\vec{v}. \end{aligned}$$

This proves that  $\vec{v}$  is an eigenvector of  $\hat{\mathbf{A}}(k)$  corresponding to the eigenvalue  $\hat{\lambda}$  for all  $k \neq 0$  and also for the case of  $\iota = 1$  for  $k = 0$ . For  $k = (0, \dots, 0)$  we note that  $\hat{\lambda}_{-1}(0) = 1$  and see that

$$\hat{\mathbf{A}}(0)\vec{v}^{(-1)}(0) = (\mathbf{C} - \mathbf{J})\vec{v}^{(-1)}(0) = -\mathbf{J}\vec{v}^{(-1)}(0) = \vec{v}^{(-1)}(0).$$

□

*Proof of Lemma 2.2.* For  $\iota \in \{1, 2, \dots, d\}$ , the vectors

$$\vec{u}^{(\iota)} = e^{ik_\iota}\vec{e}_\iota + \vec{e}_{-\iota} \quad \text{and} \quad \vec{u}^{(-\iota)} = e^{ik_\iota}\vec{e}_\iota - \vec{e}_{-\iota}$$

are eigenvectors of  $\mathbf{J}\hat{\mathbf{D}}(-k)$ , where  $\vec{u}^{(\pm\iota)}$  is associated to the eigenvalue  $\pm 1$ . For  $\iota \in \{2, 3, \dots, d\}$ , we define

$$\vec{v}^{(\iota)} = \vec{u}^{(1)} \sum_{\kappa} \vec{u}_{\kappa}^{(\iota)} - \vec{u}^{(\iota)} \sum_{\kappa} \vec{u}_{\kappa}^{(1)}, \quad \vec{v}^{(-\iota)} = \vec{u}^{(-1)} \sum_{\kappa} \vec{u}_{\kappa}^{(-\iota)} - \vec{u}^{(-\iota)} \sum_{\kappa} \vec{u}_{\kappa}^{(-1)}. \quad (2.26)$$

By construction,  $\vec{v}^{(\iota)}$  and  $\vec{v}^{(-\iota)}$  are also eigenvectors of  $\mathbf{J}\hat{\mathbf{D}}(-k)$ . For  $\mathbf{C}\hat{\mathbf{D}}(-k)$ , we compute that

$$\mathbf{C}\hat{\mathbf{D}}(-k)\vec{u}^{(\iota)} = \mathbf{C}(\vec{e}_\iota + e^{ik_\iota}\vec{e}_{-\iota}) = \sum_{\kappa} \vec{u}_{\kappa}^{(\iota)}\vec{\mathbf{1}}, \quad \mathbf{C}\hat{\mathbf{D}}(-k)\vec{u}^{(-\iota)} = \mathbf{C}(\vec{e}_\iota - e^{ik_\iota}\vec{e}_{-\iota}) = -\sum_{\kappa} \vec{u}_{\kappa}^{(-\iota)}\vec{\mathbf{1}},$$

so that

$$\mathbf{C}\hat{\mathbf{D}}(-k)\vec{v}^{(\iota)} = \mathbf{C}\hat{\mathbf{D}}(-k)\vec{v}^{(-\iota)} = 0.$$

Knowing this, it follows that

$$\hat{\mathbf{A}}(k)\vec{v}^{(\iota)} = (\mathbf{C} - \mathbf{J})\hat{\mathbf{D}}(-k)\vec{v}^{(\iota)} = -\vec{v}^{(\iota)}, \quad (2.27)$$

$$\hat{\mathbf{A}}(k)\vec{v}^{(-\iota)} = (\mathbf{C} - \mathbf{J})\hat{\mathbf{D}}(-k)\vec{v}^{(-\iota)} = \vec{v}^{(-\iota)}. \quad (2.28)$$

By (2.27),  $\vec{v}^{(\iota)}$  is an eigenvector for all  $k$ . Since the set of vectors  $(\vec{v}^{(\iota)})_{\iota=2,3,\dots,d}$  is linearly independent we know that the eigenvalue  $-1$  has geometric multiplicity  $d - 1$ .

From (2.28), we conclude the existence of  $d - 1$  linear independent eigenvalue for 1 only when  $e^{ik_\iota} \neq 1$  for all  $\iota \in \{1, 2, \dots, d\}$ . To prove that the eigenvalue 1 has geometric multiplicity  $d - 1$  for all  $k$ , we show how to choose  $d - 1$  linear independent eigenvectors when  $e^{ik_\kappa} = 1$  for a  $\kappa \in \{1, 2, \dots, d\}$ . For this, let  $S_1$  be the set of all  $\kappa \in \{1, 2, \dots, d\}$  with  $k_\kappa = 0$  and  $S_2$  the set of all  $\kappa \in \{1, 2, \dots, d\}$  with  $k_\kappa \neq 0$ . Let  $s_1$  and  $s_2$  be the number of elements in  $S_1$  and  $S_2$ . Then  $s_1 + s_2 = d$ . For all  $\iota \in S_1$ , we define  $\vec{v}^{(-\iota)} = e_\iota - e_{-\iota}$ .

If  $s_1 = d$ , then  $k = 0$ . In Lemma 2.1, we define for this case  $\hat{\lambda}_2 = 1$  and  $\vec{v}^{(-1)} = e_1 - e_{-1}$ . Then  $\{\vec{v}^{(-1)}, \vec{v}^{(-2)}, \dots, \vec{v}^{(-d)}\}$  is a set of independent eigenvectors of  $\hat{\mathbf{A}}(k)$  to the eigenvalue 1.

If  $s_1 < d$ , then let  $\rho$  be the smallest number in  $S_2$  and define

$$\vec{v}^{(-\iota)} = \vec{u}^{(-\rho)} \sum_{\kappa} \vec{u}_{\kappa}^{(-\iota)} - \vec{u}^{(-\iota)} \sum_{\kappa} \vec{u}_{\kappa}^{(-\rho)}.$$

for all  $\iota \in S_2 \setminus \{\kappa\}$ . Then it is easy to verify that the vectors  $(\vec{v}^{(-\iota)})_{\iota=2,3,\dots,d}$  are linearly independent and are eigenvectors of  $\hat{\mathbf{A}}(k)$  with eigenvalue 1. □

We use the eigensystem of the matrix  $\hat{\mathbf{A}}(k)$  to identify  $\hat{b}_n(k)$  and  $\vec{b}_n(k)$ :

**Lemma 2.3** (NBW characterization). *Let  $d \geq 2$ ,  $n \geq 1$  and  $k \in (-\pi, \pi^d)$  such that  $\hat{\lambda}_1(k) \neq \hat{\lambda}_{-1}(k)$ . Then,*

$$\hat{b}_n(k) = 2d \frac{\hat{D}(k)(\hat{\lambda}_1^n(k) - \hat{\lambda}_{-1}^n(k)) + \hat{\lambda}_{-1}^{n-1}(k) - \hat{\lambda}_1^{n-1}(k)}{\hat{\lambda}_1(k) - \hat{\lambda}_{-1}(k)}, \quad (2.29)$$

$$\vec{b}_n(k) = \frac{\hat{\lambda}_1^n(k) - \hat{\lambda}_{-1}^n(k)}{\hat{\lambda}_1(k) - \hat{\lambda}_{-1}(k)} \hat{\mathbf{D}}(k) \vec{1} - \frac{\hat{\lambda}_{-1}^{n-1}(k) - \hat{\lambda}_1^{n-1}(k)}{\hat{\lambda}_1(k) - \hat{\lambda}_{-1}(k)} \vec{1}. \quad (2.30)$$

When  $\hat{\lambda}_1(k) = \hat{\lambda}_{-1}(k)$ ,

$$\hat{b}_n(k) = 2d[(n-1)\hat{\lambda}_1(k)^{n-2} + \hat{D}(k)n\hat{\lambda}_1(k)^{n-1}], \quad \vec{b}_n(k) = [(n+1)\hat{\lambda}_1(k)^n + n\hat{\lambda}_1(k)^{n-1}\hat{\mathbf{D}}(k)]\vec{1}. \quad (2.31)$$

Clearly, one can reprove (2.20) and (2.22) using Lemma 2.3.

*Proof.* For  $\hat{\lambda}_1(k) \neq \hat{\lambda}_{-1}(k)$ , we define

$$\alpha(k) = \frac{1}{\hat{\lambda}_1(k) - \hat{\lambda}_{-1}(k)} = \frac{1}{2\sqrt{(d\hat{D}(k))^2 - (2d-1)}}.$$

We can write

$$\vec{1} = \alpha(k) \left( \hat{\lambda}_1(k)\mathbf{I} - \hat{\lambda}_{-1}(k)\mathbf{I} + \hat{\mathbf{D}}(k) - \hat{\mathbf{D}}(k) \right) \vec{1} = \alpha(k) \vec{v}^{(1)}(k) - \alpha(k) \vec{v}^{(-1)}(k).$$

Using (2.13) and the fact that  $\vec{v}^{(\pm 1)}(k)$  are eigenvectors of  $\hat{\mathbf{A}}(k)$  with eigenvalue  $\hat{\lambda}_{\pm}(k)$ , we obtain

$$\vec{b}_n(k) = \hat{\mathbf{A}}(k)^n \vec{1} = \alpha(k) \left( \hat{\lambda}_1^n(k) \vec{v}^{(1)}(k) - \hat{\lambda}_{-1}^n(k) \vec{v}^{(-1)}(k) \right), \quad (2.32)$$

which proves (2.30). Combining (2.11) and (2.32) gives

$$\hat{b}_n(k) = \alpha(k) \vec{1}^T \hat{\mathbf{D}}(-k) \left( \hat{\lambda}_1^n(k) \vec{v}^{(1)}(k) - \hat{\lambda}_{-1}^n(k) \vec{v}^{(-1)}(k) \right). \quad (2.33)$$

Inserting the definition of  $\vec{v}^{(\pm 1)}(k)$  gives (2.29). The proof of (2.31) is similar, now using that  $\hat{\mathbf{A}}(k)\vec{1} = \vec{v}^{(1)} + \hat{\lambda}_+\vec{1}$ , which implies that  $\hat{\mathbf{A}}(k)^n \vec{1} = n\hat{\lambda}_+(k)^{n-1} \vec{v}^{(1)} + \hat{\lambda}_+(k)^n \vec{1}$ .  $\square$

## 2.4 Central limit theorem

This section is devoted to the proof of a functional central limit theorem for the NBW. The uniform measures on  $n$ -step NBWs form a consistent family of measures, so that there is a unique law that describes them as a process. We let  $\omega = (\omega_r)_{r \geq 0}$  be distributed according to this law. For a NBW  $\omega$  and  $t \geq 0$ , we define

$$X_n(t) = \frac{\omega_{\lfloor nt \rfloor}}{\sqrt{n}}, \quad (2.34)$$

where  $\lfloor x \rfloor$  denotes the integer part of  $x \in \mathbb{R}$ .

**Theorem 2.4** (Functional central limit theorem). *The processes  $(X_n(t))_{t \geq 0}$  converge weakly to  $(B(t))_{t \geq 0}$ , where  $(B(t))_{t \geq 0}$  is a Brownian motion with covariance matrix  $\mathbf{I}/(d-1)$ .*

Our proof of Theorem 2.4 is organized as follows. We use Lemma 2.3 to prove a CLT for the endpoint in Lemma 2.5. We then prove the convergence of the finite-dimensional distributions to the Gaussian distribution (see Lemma 2.6), followed by a proof of tightness (see Lemma 2.8). This implies Theorem 2.4, see e.g. [5, Theorem 15.6]. We now fill in the details.

**Lemma 2.5** (Central limit theorem). *Let  $d \geq 2$  and  $k \in [-\pi, \pi]^d$ . Then,*

$$\lim_{n \rightarrow \infty} \mathbb{E}[e^{ik \cdot \omega_n / \sqrt{n}}] = e^{-\|k\|_2^2 / (2d-2)}, \quad (2.35)$$

where  $\|k\|_2^2 = \sum_{i=1}^d k_i^2$  is the Euclidean norm of  $k$ .

Lemma 2.5 implies that the distribution of the endpoint of an  $n$ -step NBW converges in distribution to a normal distribution with mean zero and covariance matrix  $(d-1)^{-1}\mathbf{I}$ .

*Proof.* We can rewrite the expectation as

$$\mathbb{E}[e^{ik \cdot \omega_n / \sqrt{n}}] = \sum_{x \in \mathbb{Z}} \frac{b_n(x)}{\hat{b}_n(0)} e^{ik \cdot x / \sqrt{n}} = \frac{\hat{b}_n(k / \sqrt{n})}{2d(2d-1)^n}. \quad (2.36)$$

As  $n \rightarrow \infty$ , we can assume without loss of generality that  $k$  is small and therefore that  $\hat{\lambda}_1(k) > \hat{\lambda}_{-1}(k)$ . Then we can use (2.29) to compute the limit of (2.36). For  $\iota = 1, -1$ , we compute

$$\lim_{n \rightarrow \infty} \frac{\hat{\lambda}_\iota^{n-1}(k / \sqrt{n})}{(2d-1)^{n-1}} \alpha(k / \sqrt{n}) \frac{\vec{1}^T \hat{\mathbf{D}}(-k / \sqrt{n}) \vec{v}^{(\iota)}(k / \sqrt{n})}{2d}. \quad (2.37)$$

We first consider the case  $\iota = 1$ . The coefficient  $\alpha(k)$  as well as  $\vec{v}^{(1)}(k)$  are continuous in a neighborhood of  $\vec{0}$ , so we can directly compute

$$\lim_{n \rightarrow \infty} \alpha(k) \frac{\vec{1}^T \hat{\mathbf{D}}(-k / \sqrt{n}) \vec{v}^{(1)}(k / \sqrt{n})}{2d} = \alpha(0) \frac{\vec{1}^T \hat{\mathbf{D}}(0) \vec{v}^{(1)}(0)}{2d} = 1.$$

Further,  $\hat{\lambda}_1(k)$  is differentiable in  $k$ , so we can Taylor expand  $\hat{\lambda}_1(k)$  at 0 to obtain

$$\hat{\lambda}_1(k) = 2d - 1 - \frac{2d-1}{2d-2} \|k\|_2^2 + O(\|k\|_2^4).$$

Therefore,

$$\lim_{n \rightarrow \infty} \frac{\hat{\lambda}_1^{n-1}(k / \sqrt{n})}{(2d-1)^{n-1}} = \lim_{n \rightarrow \infty} \left( 1 - \frac{1}{2d-2} \frac{\|k\|_2^2}{n} + O\left(\frac{\|k\|_2^4}{n^2}\right) \right)^{n-1} = e^{-\|k\|_2^2 / (2d-2)}, \quad (2.38)$$

so that the limit (2.37) for  $\iota = 1$  is given by  $e^{-\frac{1}{2d-2} \|k\|_2^2}$ .

We next consider the case  $\iota = -1$ , for which we use that  $k \mapsto \hat{\lambda}_{-1}(k)$  is continuous and  $\hat{\lambda}_{-1}(0) = 1$ . Therefore, for  $d \geq 2$ ,

$$\lim_{n \rightarrow \infty} \frac{\hat{\lambda}_{-1}^{n-1}(k / \sqrt{n})}{(2d-1)^{n-1}} = 0. \quad (2.39)$$

The second factor in (2.37) can easily be bounded uniformly for small  $k$ , so that for  $\iota = -1$  the limit of (2.37) is zero.  $\square$

**Lemma 2.6** (Convergence of finite-dimensional distributions). *For  $d \geq 2$  and  $N > 0$ , let  $0 = t_0 < t_1 < t_2 < \dots < t_N = 1$  and  $k^{(r)} \in (-\pi, \pi]^d$  for  $r = 1, \dots, N$ . Then,*

$$\lim_{n \rightarrow \infty} \mathbb{E}[e^{i(\sum_{r=1}^N k^{(r)} (\omega_{\lfloor t_r n \rfloor} - \omega_{\lfloor t_{r-1} n \rfloor}) / \sqrt{n})}] = e^{-\sum_{r=1}^N \|k^{(r)}\|_2^2 (t_r - t_{r-1}) / (2d-2)}. \quad (2.40)$$

*Proof.* As we take the limit  $n \rightarrow \infty$ , without loss of generality we can assume that  $\eta_r(n) := \lfloor t_r n \rfloor - \lfloor t_{r-1} n \rfloor \geq 1$ ,  $t_r > t_{r-1}$  and each  $k_n^{(r)} := k^{(r)}/\sqrt{n}$  is so small that  $\hat{\lambda}_1(k) > \hat{\lambda}_{-1}(k)$  for  $r = 1, \dots, N$  and  $n \in \mathbb{N}$ . Let  $\mathcal{W}_n$  be the set of all  $n$ -step NBW. For any function  $f: \mathbb{Z}^{d \times N} \mapsto \mathbb{C}$ , we know that

$$\begin{aligned} \mathbb{E}[f(\omega_{\lfloor t_1 n \rfloor}, \dots, \omega_{\lfloor t_N n \rfloor})] &= \frac{1}{\hat{b}_n(0)} \sum_{\omega \in \mathcal{W}_n} f(\omega_{\lfloor t_1 n \rfloor}, \dots, \omega_{\lfloor t_N n \rfloor}) \\ &= \frac{1}{\hat{b}_n(0)} \sum_{\omega \in \mathcal{W}_n} \sum_{x_1, \dots, x_N \in \mathbb{Z}} f(x_1, \dots, x_N) \prod_{i=1, \dots, N} \delta_{x_i, \lfloor \omega_{t_i n} \rfloor}. \end{aligned} \quad (2.41)$$

Let  $b_n^{\iota, \kappa}(x)$  be the number of  $n$ -step NBW  $\omega$  with  $\omega_1 \neq e_\iota$ ,  $\omega_{n-1} = x + e_\kappa$  and  $\omega_n = x$ . We define the matrix  $\hat{\mathbf{B}}_n(k)$  with entries  $(\hat{\mathbf{B}}_n(k))_{\iota, \kappa} = \hat{b}_n^{\iota, \kappa}(k)$ . By a relation similar to (2.10), we conclude that  $\hat{\mathbf{B}}_n(k) = \hat{\mathbf{A}}(k)^n$ . We fix  $N$  points  $x_1, \dots, x_N \in \mathbb{Z}$ , then the number of NBW  $\omega$  with  $\omega_{\lfloor t_i n \rfloor} = x_i$  for all  $i = 1, 2, \dots, N$  is given by

$$\sum_{\substack{\iota_1, \dots, \iota_N \\ \text{in } \{\pm 1, \dots, \pm d\}}} b_{\eta_1(n)-1}^{\iota_0, \iota_1}(x_1 + e_{\iota_0}) \prod_{r=2, \dots, N} b_{\eta_r(n)}^{\iota_{r-1}, \iota_r}(x_r - x_{r-1}). \quad (2.42)$$

We insert  $f(x_1, \dots, x_N) = e^{i \sum_r k_n^{(r)}(x_r - x_{r-1})}$  and obtain

$$\begin{aligned} \mathbb{E}[e^{i \sum_r k_n^{(r)}(\omega_{n_r} - \omega_{n_{r-1}})}] &= \frac{1}{\hat{b}_n(0)} \sum_{\substack{\iota_1, \dots, \iota_N \\ \text{in } \{\pm 1, \dots, \pm d\}}} \hat{b}_{\eta_1(n)-1}^{\iota_1}(k_n^{(1)}) e^{-i k_{\iota_1}^{(1)}/\sqrt{n}} \prod_{r=2}^N \hat{b}_{\eta_r(n)}^{\iota_{r-1}, \iota_r}(k_n^{(r)}) \\ &= \frac{1}{\hat{b}_n(0)} \vec{1}^T \hat{\mathbf{D}}(-k_n^{(1)}) \hat{\mathbf{A}}(k_n^{(1)})^{\eta_1(n)-1} \prod_{r=2}^N \hat{\mathbf{B}}_{\eta_r(n)}(k_n^{(r)}) \vec{1} \\ &= \frac{1}{\hat{b}_n(0)} \vec{1}^T \hat{\mathbf{D}}(-k_n^{(1)}) \hat{\mathbf{A}}(k_n^{(1)})^{\eta_1(n)-1} \prod_{r=2}^N \hat{\mathbf{A}}(k_n^{(r)})^{\eta_r(n)} \vec{1}. \end{aligned} \quad (2.43)$$

To proceed, we define, for  $\tau \in \{1, 2, \dots, N\}$ ,

$$\vec{h}_n(\tau) = \frac{1}{\hat{b}_{\lfloor t_\tau n \rfloor}(0)} \vec{1}^T \hat{\mathbf{D}}(-k_n^{(1)}) \hat{\mathbf{A}}(k_n^{(1)})^{\eta_1(n)-1} \dots \hat{\mathbf{A}}(k_n^{(\tau)})^{\eta_\tau(n)}, \quad \vec{h}_n(0) = \frac{1}{2d} \vec{1}^T \hat{\mathbf{D}}(-k_n^{(1)}).$$

As  $\sum_{r=1}^\tau \eta_r(n) \leq t_\tau n$ , by construction of  $\vec{h}_n(\tau)$ , for all  $\vec{v} \in \mathbb{C}^{2d}$

$$\begin{aligned} |\vec{h}_n(\tau) \vec{v}| &\leq \mathbb{E} \left[ e^{i(\sum_{r=1}^\tau k^{(r)}(\omega_{\lfloor t_r n \rfloor} - \omega_{\lfloor t_{r-1} n \rfloor})/\sqrt{n})} \sum_{\iota} \mathbb{1}_{\{\omega_{t_\tau n} - \omega_{t_\tau n-1} = e_\iota\}} |\vec{v}_\iota| \right] \\ &\leq \mathbb{E} \left[ \sum_{\iota} \mathbb{1}_{\{\omega_{t_\tau n} - \omega_{t_\tau n-1} = e_\iota\}} |\vec{v}_\iota| \right] \leq \|\vec{v}\|_\infty, \end{aligned} \quad (2.44)$$



where we write  $\|v\|_\infty := \max_l |v_l|$ . Therefore, we can rewrite

$$\begin{aligned}
(2.43) &= \vec{h}_n(N) \vec{1} = \vec{h}_n(N-1) \left( \frac{\hat{\mathbf{A}}(k_n^{(N)})}{2d-1} \right)^{\eta_N(n)} \frac{\vec{v}^{(1)}(k_n^{(N)})}{2d-2} + \vec{h}_n(N) \left( \vec{1} - \frac{\vec{v}^{(1)}(k_n^{(N)})}{2d-2} \right) \\
&= \vec{h}_n(N-1) \left( \frac{\hat{\lambda}_1(k_n^{(N)})}{2d-1} \right)^{\eta_N(n)} \frac{\vec{v}^{(1)}(k_n^{(N)})}{2d-2} + \vec{h}_n(N) \left( \vec{1} - \frac{\vec{v}^{(1)}(k_n^{(N)})}{2d-2} \right) \\
&= \vec{h}_n(0) \frac{\vec{v}^{(1)}(k_n^{(1)})}{2d-2} \left( \frac{2d-1}{\hat{\lambda}_1(k_n^{(1)})} \right) \prod_{r=1}^N \left( \frac{\hat{\lambda}_1(k_n^{(r)})}{2d-1} \right)^{\eta_r(n)} \tag{2.45}
\end{aligned}$$

$$+ \sum_{r=1}^{N-1} \frac{\vec{h}_n(r)}{2d-2} (\vec{v}^{(1)}(k_n^{(r+1)}) - \vec{v}^{(1)}(k_n^{(r)})) \prod_{i=r+1}^N \left( \frac{\hat{\lambda}_1(k_n^{(i)})}{2d-1} \right)^{\eta_i(n)} \tag{2.46}$$

$$+ \vec{h}_n(N) \left( \vec{1} - \frac{\vec{v}^{(1)}(k_n^{(N)})}{2d-2} \right). \tag{2.47}$$

To compute the limit of (2.43), we show that (2.46) and (2.47) converge to zero and identify the limit of (2.45). From (2.44), it follows that we can bound (2.47) by  $\vec{1} - \vec{v}^{(1)}(k_n^{(N)})/(2d-2)$  and we know that  $\vec{v}^{(1)}(k_n^{(N)}) \xrightarrow{n \rightarrow \infty} (2d-2)\vec{1}$ . Thereby we conclude that (2.47) converges to 0.

To compute the limit of (2.46) we note that  $\hat{\lambda}_1(k) \leq 2d-1$  and that  $\vec{v}^{(r+1)}(k_n^{(r+1)}) - \vec{v}^{(r)}(k_n^{(r)}) \xrightarrow{n \rightarrow \infty} \vec{0}$ , so that

$$\begin{aligned}
\lim_{n \rightarrow \infty} |(2.46)| &\leq \lim_{n \rightarrow \infty} \left| \sum_{r=1}^{N-1} \frac{\vec{h}_n(r)}{2d-2} (\vec{v}^{(1)}(k_n^{(1)}) - \vec{v}^{(1)}(k_n^{(1)})) \right| \\
&\stackrel{(2.44)}{\leq} \lim_{n \rightarrow \infty} \frac{1}{2d-2} \sum_{r=1}^{N-1} \|\vec{v}^{(1)}(k_n^{(1)}) - \vec{v}^{(1)}(k_n^{(1)})\|_\infty = 0.
\end{aligned}$$

To compute the limit of (2.45), we use that

$$\lim_{n \rightarrow \infty} \vec{h}_n(0) \frac{\vec{v}^{(1)}(k_n^{(1)})}{2d-2} \frac{2d-1}{\hat{\lambda}_1(k_n^{(1)})} = \lim_{n \rightarrow \infty} \frac{\vec{1}^T \hat{\mathbf{D}}(-k) (\hat{\lambda}_1(k_n^{(1)}) \vec{1} + \hat{\mathbf{D}}(k_n^{(1)}) \vec{1})}{2d(2d-2)} \frac{2d-1}{\hat{\lambda}_1(k_n^{(1)})} = 1,$$

and, for each  $r \in \{1, \dots, N\}$ ,

$$\lim_{n \rightarrow \infty} \left( \frac{\hat{\lambda}_1(k_n^{(r)})}{2d-1} \right)^{\eta_r(n)} \stackrel{(2.38)}{=} e^{-\|k^{(i)}\|_2^2(t_i - t_{i-1})/(2d-2)}.$$

Combining this yields that the limit of (2.45) is the right-hand side of (2.40), which completes the proof of Lemma 2.6.  $\square$

To prove tightness we make use of the following lemma, which computes the second moment of the end-point of NBW. The leading order in this result was alternatively proved in [9, (5.3.11)] using residues.

**Lemma 2.7** (The second moment). *For  $d \geq 2$  and  $n \in \mathbb{N}$ ,*

$$\mathbb{E}[\|\omega_n\|_2^2] = \frac{d}{d-1}n + \frac{4d-1}{2(d-1)^2} + \frac{d}{2(d-1)^2(2d-1)^{n-2}}.$$

*Proof.* We define the differential operator  $\nabla^2 := \sum_{i=1}^d \left( \frac{\partial}{\partial k_i} \right)^2$  and see that

$$\mathbb{E}[\|\omega_n\|_2^2] = \frac{1}{\hat{b}_n(0)} \sum_{x \in \mathbb{Z}} b_n(x) \|x\|_2^2 = -\frac{1}{\hat{b}_n(0)} \nabla^2 \hat{b}_n(0).$$

To compute the second derivative in a neighborhood of the origin, we recall (2.29). By Lemma 2.3, for  $k$  such that  $\hat{\lambda}_+(k) \neq \hat{\lambda}_-(k)$ ,

$$\hat{b}_n(k) = \sum_{\sigma \in \{-1,1\}} d\nu \frac{\hat{\lambda}_\sigma^n(k) \hat{D}(k) - \hat{\lambda}_\sigma^{n-1}(k)}{\sqrt{(d\hat{D}(k))^2 - (2d-1)}} \equiv \sum_{\sigma \in \{-1,1\}} \hat{g}_\sigma(k). \quad (2.48)$$

Since,  $\hat{\lambda}_+(k) \neq \hat{\lambda}_-(k)$  for  $k$  small, (2.48) holds in particular for  $k$  small. A straightforward computation gives

$$\begin{aligned} \nabla^2 g_\sigma(k) \Big|_{k=0} &= d\sigma \left( \nabla^2 \hat{\lambda}_\sigma(0) \right) \frac{n\hat{\lambda}_\sigma^{n-1}(0) \hat{D}(0) - (n-1)\hat{\lambda}_\sigma^{n-2}(0)}{\sqrt{(d\hat{D}(0))^2 - (2d-1)}} \\ &\quad + d\sigma \left( \nabla^2 \hat{D}(0) \right) \left( \frac{\hat{\lambda}_\sigma^n(0)}{\sqrt{(d\hat{D}(0))^2 - (2d-1)}} - d^2 \hat{D}(0) \frac{\hat{\lambda}_\sigma^n(0) \hat{D}(0) - \hat{\lambda}_\sigma^{n-1}(0)}{\left( \sqrt{(d\hat{D}(0))^2 - (2d-1)} \right)^3} \right) \\ &= -d\sigma \frac{d((1+\sigma)d-1)}{(d-1)} \frac{n\hat{\lambda}_\sigma^{n-1}(0) - (n-1)\hat{\lambda}_\sigma^{n-2}(0)}{d-1} \\ &\quad + \frac{d\sigma}{d-1} \hat{\lambda}_\sigma^{n-1}(0) \left( \hat{\lambda}_\sigma(0) - d^2 \frac{\hat{\lambda}_\sigma(0) - 1}{(d-1)^2} \right), \end{aligned}$$

so that

$$\begin{aligned} \nabla^2 g_1(k) \Big|_{k=0} &= -\frac{d}{d-1} \frac{(2d-1)^{n-1}}{d-1} [2d(d-1)n + (4d-1)] \\ \nabla^2 g_{-1}(k) \Big|_{k=0} &= \frac{d}{d-1} \left[ \frac{d(-1)}{(d-1)} - 1 \right] = -d \frac{2d-1}{(d-1)^2}. \end{aligned}$$

We arrive at

$$\begin{aligned} -\frac{1}{\hat{b}_n(0)} \nabla^2 \hat{b}_n(0) &= \frac{1}{2d(2d-1)^{n-1}} \sum_{\sigma \in \{-1,1\}} -\nabla^2 g_\sigma(k) \Big|_{k=0} \\ &= \frac{d}{d-1} n + \frac{4d-1}{2(d-1)^2} + \frac{d}{2(d-1)^2(2d-1)^{n-2}}. \end{aligned}$$

□

**Lemma 2.8** (Tightness). *For  $d \geq 2$  and  $0 \leq t_1 < t_2 < t_3 \leq 1$ , there exists a  $K > 0$ , such that, for all  $n \geq 1$ ,*

$$\mathbb{E}[\|X_n(t_2) - X_n(t_1)\|_2^2 \|X_n(t_3) - X_n(t_2)\|_2^2] \leq K(t_2 - t_1)(t_3 - t_2).$$

*Proof.* We use the same notation as in the proof of Lemma 2.6. In (2.42), we have seen how to describe the number of NBWs that visit a number of fixed points. This time we forget about the non-backtracking

constraint between two subsequent NBWs to upper bound

$$\begin{aligned}
& \mathbb{E}[\|\omega_{[t_2n]} - \omega_{[t_1n]}\|_2^2 \|\omega_{[t_3n]} - \omega_{[t_2n]}\|_2^2] \\
& \leq \frac{1}{\hat{b}_n(0)} \sum_{x_1, x_2, x_3, x_4 \in \mathbb{Z}^d} \|x_2 - x_1\|_2^2 \|x_3 - x_2\|_2^2 b_{\eta_1(n)}(x_1) \prod_{r=2}^4 b_{\eta_r(n)}(x_r - x_{r-1}) \\
& = \frac{1}{\hat{b}_n(0)} \sum_{x_1, x_2, x_3, x_4 \in \mathbb{Z}^d} \|x_2\|_2^2 \|x_3\|_2^2 \prod_{r=1}^4 b_{\eta_r(n)}(x_r) \\
& = \frac{\hat{b}_{\eta_1(n)} \hat{b}_{\eta_4(n)}}{b_n} \sum_{x_2 \in \mathbb{Z}} b_{\eta_2(n)}(x_2) \|x_2\|_2^2 \sum_{x_3 \in \mathbb{Z}} b_{\eta_3(n)}(x_3) \|x_3\|_2^2 \\
& = \left( \frac{2d-1}{2d} \right)^3 \mathbb{E}[\|\omega_{\eta_2(n)}\|_2^2] \mathbb{E}[\|\omega_{\eta_3(n)}\|_2^2].
\end{aligned}$$

Applying Lemma 2.7 completes the proof.  $\square$

## 2.5 Extension to non-nearest-neighbor setting

In this section, we extend the analysis of NBW on  $\mathbb{Z}^d$  to other bond sets. We start by introducing the bond sets that we consider. We let  $\mathbb{B} \subset \mathbb{Z}^d \times \mathbb{Z}^d$  be a translation invariant collection of bonds. Let  $\mathbb{V}_0 = \{x: \{0, x\} \in \mathbb{B}\}$  denote the set of endpoints of bonds containing the origin and write  $m = |\mathbb{V}_0|$ . We assume that  $0 \notin \mathbb{V}_0$ , and that  $\mathbb{V}_0$  is symmetric, i.e.,  $-x \in \mathbb{V}_0$  for every  $x \in \mathbb{V}_0$ . Thus,  $m$  is even. We define the simple random walk step distribution by

$$D(x) = \frac{1}{m} \mathbb{1}_{\{x \in \mathbb{V}_0\}}. \quad (2.49)$$

Define the matrices  $\mathbf{C}, \mathbf{J} \in \mathbb{C}^{m \times m}$  by  $(\mathbf{C})_{x,y} = 1$  and  $(\mathbf{J})_{x,y} = \delta_{x,-y}$ , and let the diagonal matrix  $\hat{\mathbf{D}}(k)$  have entries  $(\hat{\mathbf{D}}(k))_{x,x} = e^{ik \cdot x}$ , where  $x, y \in \mathbb{B}_0$ . Then, we define the matrix  $\hat{\mathbf{A}}(k)$  of size  $m \times m$  by

$$\hat{\mathbf{A}}(k) = (\mathbf{C} - \mathbf{J})\hat{\mathbf{D}}(-k). \quad (2.50)$$

With this definition at hand, we see that (2.11)–(2.13) remain to hold. As a result, also Lemmas 2.1–2.2, whose proof only depends on (2.11)–(2.13), continue to hold when we replace each occurrence of  $2d$  by  $m$ . Since the proof of Lemma 2.3, in turn, only depends on Lemmas 2.1–2.2, also it extends to this setting, so that, for example

$$\hat{\lambda}_{\pm}(k) = F_{\pm}(\hat{D}(k); m), \quad \text{where} \quad F_{\pm}(x; m) = \frac{1}{2} \left( mx \pm \sqrt{(mx)^2 - 4(m-1)} \right), \quad (2.51)$$

and, when  $\hat{\lambda}_1(k) \neq \hat{\lambda}_{-1}(k)$ ,

$$\hat{b}_n(k) = \frac{m}{2} \frac{\hat{D}(k)(\hat{\lambda}_1^n(k) - \hat{\lambda}_{-1}^n(k)) + (\hat{\lambda}_{-1}^{n-1}(k) - \hat{\lambda}_1^{n-1}(k))}{\hat{\lambda}_1(k) - \hat{\lambda}_{-1}(k)}. \quad (2.52)$$

Naturally, Theorem 2.4 needs to be adapted, and now reads that the processes  $(X_n(t))_{t \geq 0}$  converge weakly to a Brownian motion with covariance matrix  $\mathbf{M}$  of size  $d \times d$ , where, for  $\iota, \kappa \in \{\pm 1, \dots, \pm d\}$ , we define

$$\mathbf{M}_{\iota, \kappa} = \frac{\partial^2 \hat{\lambda}_1(k)}{\partial k_{\iota} \partial k_{\kappa}} \Big|_{k=0} \hat{\lambda}_1(0)^{-1}. \quad (2.53)$$

We next compute  $\mathbf{M}$  explicitly in terms of the covariance matrix of the transition kernel  $D$ . We compute that  $F_+(1; m) = m/2 + (m-2)/2 = m-1$ , and

$$F'_+(x; m) = m/2 + \frac{m^2 x}{2\sqrt{(mx)^2 - 4(m-1)}}, \quad \text{so that} \quad F'_+(1; m) = m(m-1)/(m-2). \quad (2.54)$$

By symmetry, the odd derivatives of  $\hat{D}(k)$  are zero, so that a Taylor expansion yields

$$\hat{D}(k) = 1 - \frac{1}{2}k^T \mathbf{H}k + O(\|k\|_2^4), \quad (2.55)$$

where, for  $\iota, \kappa \in \{1, \dots, d\}$ ,

$$\mathbf{H}_{\iota, \kappa} = \sum_x x_\iota x_\kappa D(x) \quad (2.56)$$

denotes the covariance matrix of SRW. As a result,

$$\mathbf{M} = \mathbf{H} \frac{F'_+(1; m)}{F_+(1; m)} = \mathbf{H}m/(m-2). \quad (2.57)$$

In the nearest-neighbor case,  $m = 2d$  and  $\mathbf{H} = \mathbf{I}/d$ , so that we retrieve the result in Theorem 2.4.

### 3 NBW on tori

In this section, we extend the results in Section 2 to NBWs on tori. In Section 3.1 and 3.2, we investigate NBW on a torus of width  $r \geq 2$ , and in Section 3.3 we investigate NBW on the hypercube, for which  $r = 2$ . The study of random walks on various finite transitive graphs has attracted considerable attention. See e.g., [8] for a recent book on the subject, and [2] for a book in preparation. Here we restrict ourselves to NBWs on tori.

#### 3.1 Setting

For  $d \geq 2$  and  $r \geq 3$ , we denote by  $\mathbb{T} = \mathbb{T}_{r,d} = (\mathbb{Z}/r\mathbb{Z})^d$  the discrete  $d$ -dimensional torus with side length  $r$ . The torus has periodic boundaries, i.e., we identify two points  $x, y \in \mathbb{T}_{r,d}$  if  $x_i \bmod r = y_i \bmod r$  for all  $i = 1, \dots, d$  where mod denote the modulus. We define the Fourier dual torus of  $\mathbb{T}$  as

$$\mathbb{T}_{r,d}^* := \frac{2\pi}{r} \left\{ -\left\lfloor \frac{r-1}{2} \right\rfloor, \dots, \left\lfloor \frac{r-1}{2} \right\rfloor \right\}^d, \quad (3.1)$$

so that each component of  $k \in \mathbb{T}_{r,d}^*$  is between  $-\pi$  and  $\pi$ . The Fourier transform of  $f: \mathbb{T}_{r,d} \rightarrow \mathbb{C}$  is defined by

$$\hat{f}(k) = \sum_{x \in \mathbb{T}_{r,d}} f(x) e^{ik \cdot x}, \quad k \in \mathbb{T}_{r,d}^*.$$

As in Section 2, we define an  $n$ -step random walk on  $\mathbb{T}_{r,d}$  to be an ordered tuple  $\omega = (\omega_0, \dots, \omega_n)$ , with  $\omega_i \in \mathbb{T}_{r,d}$  and  $\omega_i - \omega_{i+1} \in \mathbb{V}_0$ , where we recall that  $\mathbb{V}_0 = \{x: \{0, x\} \in \mathbb{B}\}$ , and  $\mathbb{B}$  is the translationally invariant bond set on which our random walks moves. We always assume that  $0 \notin \mathbb{V}_0$ , and that  $\mathbb{V}_0$  is symmetric, i.e., if  $x \in \mathbb{V}_0$ , then also  $-x \in \mathbb{V}_0$ . Further, we always assume that  $\omega_0 = (0, \dots, 0)$ . The simple random walk step distribution is given by

$$D(x) = \frac{1}{m} \mathbb{1}_{\{x \in \mathbb{V}_0\}} \quad \text{and} \quad \hat{D}(k) = \frac{1}{m} \sum_{x \in \mathbb{T}_{r,d}} e^{ik \cdot x}. \quad (3.2)$$

If an  $n$ -step random walk on  $\mathbb{T}_{r,d}$  additionally satisfies  $\omega_i \neq \omega_{i-2}$ , then we call the walk a non-backtracking walk (NBW) on  $\mathbb{T}_{r,d}$ . Let  $b_n(x)$  be the number of  $n$ -step NBWs with  $\omega_n = x$ . Further, let  $b'_n(x)$  be the number of  $n$ -step NBWs  $\omega$  with  $\omega_n = x$  and  $\omega_1 \neq e_\iota$ .

In this setting, we can express  $b_n(x)$  for NBW on  $\mathbb{T}_{r,d}$  in terms of NBW on  $\mathbb{Z}^d$ . Indeed, identify  $\mathbb{T}_{r,d}$  with  $\{0, \dots, r-1\}^d \subset \mathbb{Z}^d$ , and also identify  $\mathbb{V}_0 = \{x: \{0, x\} \in \mathbb{B}\}$  as a subset of  $\{0, \dots, r-1\}^d \subset \mathbb{Z}^d$ .

Define  $\mathbb{V}_0^{\mathbb{Z}} = \mathbb{V}_0 \cup (-\mathbb{V}_0)$  (which, by construction, are disjoint subsets of  $\mathbb{Z}^d$ ), and define the random walk step distribution  $D_{\mathbb{Z}}(x)$  by the uniform distribution on  $\mathbb{V}_0^{\mathbb{Z}}$ . Then, for  $x \in \mathbb{T}_{r,d}$ ,

$$b_n(x) = \sum_{y: y \stackrel{r}{\sim} x} b_n^{\mathbb{Z}}(y), \quad (3.3)$$

where, for  $y \in \mathbb{Z}^d$  and  $x \in \mathbb{T}_{r,d}$ , we say that  $x \stackrel{r}{\sim} y$  when  $x = y \pmod{r}$ , and  $b_n^{\mathbb{Z}}(y)$  denotes the number of  $n$ -step NBWs on  $\mathbb{Z}^d$  with step distribution  $D_{\mathbb{Z}}(x)$ . As a result,  $\hat{b}_n(k)$  for NBW on  $\mathbb{T}_{r,d}$  is equal to  $\hat{b}_n^{\mathbb{Z}}(k)$  for every  $k \in \mathbb{T}_{r,d}^*$  as defined in (3.1). Therefore, we can use most results for NBW on  $\mathbb{Z}^d$  to study NBW on  $\mathbb{T}_{r,d}$ . We define the probability mass function of the endpoint of an  $n$ -step NBW by

$$p_n(x) = \frac{b_n(x)}{\sum_y b_n(y)} = \frac{b_n(x)}{m(m-1)^{n-1}}. \quad (3.4)$$

In order to study the asymptotic behavior of NBW, we investigate  $\hat{\lambda}_{\pm}(k)$  for  $k \neq 0$ . Our main result in this section is the following theorem:

**Theorem 3.1** (Pointwise bound on  $\hat{b}_n(k)$ ). *Let  $\mathbb{V}_0$  be symmetric and satisfy  $0 \notin \mathbb{V}_0$ . Then, for  $n \in \mathbb{N}$ , NBW with steps in  $\mathbb{V}_0$  satisfies*

$$|\hat{p}_n(k)| \leq (1/\sqrt{m-1} \vee |\hat{D}(k)|)^{n-1}. \quad (3.5)$$

To prove Theorem 3.1, we start by investigating  $\hat{\lambda}_{\pm}(k)$  for  $k \neq 0$ . For this, we use Lemma 2.1 to note that

$$\hat{\lambda}_{\pm}(k) = F_{\pm}(\hat{D}(k); m), \quad \text{where} \quad F_{\pm}(x; m) = \frac{1}{2} \left( mx \pm \sqrt{(mx)^2 - 4(m-1)} \right), \quad (3.6)$$

and where  $m$  denotes the degree of our graph. We bound  $\hat{\lambda}_{\pm}(k)$  in the following lemma:

**Lemma 3.2** (Bounds on  $\hat{\lambda}_{\pm}(k)$ ). *For any  $k \in \mathbb{T}_{r,d}^*$ ,*

$$|\hat{\lambda}_{+}(k)| \begin{cases} = \sqrt{m-1} & \text{when } (m\hat{D}(k))^2 - 4(m-1) \leq 0; \\ \leq (m-1)[1 - (1 - \hat{D}(k))m/(m-2)] & \text{when } \hat{D}(k) \geq 0, (m\hat{D}(k))^2 - 4(m-1) > 0; \\ \leq 1 & \text{when } \hat{D}(k) \leq 0, (m\hat{D}(k))^2 - 4(m-1) > 0. \end{cases} \quad (3.7)$$

*Proof.* The function  $x \mapsto F_{\pm}(x; m)$  is real when  $(mx)^2 - 4(m-1) \geq 0$ , and complex when  $(mx)^2 - 4(m-1) < 0$ . When  $(mx)^2 - 4(m-1) < 0$ ,

$$|F_{\pm}(x; m)|^2 = m-1, \quad (3.8)$$

so that  $|F_{\pm}(x; m)| = \sqrt{m-1}$ .

When  $(mx)^2 - 4(m-1) \geq 0$ , by the symmetry  $F_{+}(x; m) = F_{-}(-x; m)$ , we only need to investigate  $x \in [0, 1]$ . We start with  $F_{-}(x; m)$ , which clearly satisfies  $F_{-}(x; m) \geq 0$ . Further, we can compute that  $F_{-}(1; m) = 1$ , and

$$F'_{-}(x; m) = m/2 - \frac{m^2 x}{2\sqrt{(mx)^2 - 4(m-1)}} = \frac{m}{2} \left[ 1 - \frac{mx}{\sqrt{(mx)^2 - 4(m-1)}} \right] < 0, \quad (3.9)$$

so that  $F_{-}(x; m) \leq 1$  for all  $x \in [0, 1]$  for which  $(mx)^2 - 4(m-1) > 0$ .

To bound  $F_{+}(x; m)$ , we use (2.54) as well as

$$\begin{aligned} F''_{+}(x; m) &= \frac{m^2}{2\sqrt{(mx)^2 - 4(m-1)}} - \frac{m^4 x^2}{2((mx)^2 - 4(m-1))^{3/2}} \\ &= \frac{m^2}{2((mx)^2 - 4(m-1))^{3/2}} \{((mx)^2 - 4(m-1)) - (mx)^2\} \\ &= -\frac{2m^2(m-1)}{((mx)^2 - 4(m-1))^{3/2}} < 0. \end{aligned} \quad (3.10)$$

As a result, a Taylor expansion yields

$$\begin{aligned} F_+(x; m) &\leq F_+(1; m) + (x-1)F'_+(1; m) = (m-1) + (x-1)m(m-1)/(m-2) \\ &= (m-1)[1 - (1-x)m/(m-2)]. \end{aligned} \quad (3.11)$$

□

*Proof of Theorem 3.1.* By (2.11) and (2.13),

$$|\hat{b}_n(k)| \leq \|\hat{\mathbf{D}}(-k)\vec{1}\|_2 \|\vec{b}_{n-1}(k)\|_2 \leq \|\vec{1}\|_2 \|\hat{\mathbf{A}}(k)\|_{\text{OP}}^{n-1} \|\vec{1}\|_2, \quad (3.12)$$

where we write  $\|\mathbf{M}\|_{\text{OP}} = \sup \|\mathbf{M}x\|_2 / \|x\|_2$  for the operator norm of the matrix  $\mathbf{M}$ . We next use that  $\hat{\mathbf{A}}(k)$  has eigenvalues  $\hat{\lambda}_+(k)$ ,  $\hat{\lambda}_-(k)$  and  $\pm 1$  by Lemmas 2.1-2.2, so that

$$\|\hat{\mathbf{A}}(k)\|_{\text{OP}} = |\hat{\lambda}_+(k)| \vee |\hat{\lambda}_-(k)| \vee 1, \quad (3.13)$$

where we use that for finite-dimensional matrices, the operator norm is equal to the maximal eigenvalue, and for  $x, y \in \mathbb{R}$ , we write  $(x \vee y) = \max\{x, y\}$ . Thus, we arrive at

$$|\hat{b}_n(k)| \leq m(|\hat{\lambda}_+(k)| \vee |\hat{\lambda}_-(k)| \vee 1)^{n-1}. \quad (3.14)$$

By Lemma 3.2, and since  $m \geq 2$  so that  $\sqrt{m-1} \geq 1$ ,

$$\frac{|\hat{\lambda}_{\pm}(k)|}{m-1} \leq (m-1)^{-1/2} \vee (1 - [1 - \hat{D}(k)] \frac{m}{m-2}) \leq (m-1)^{-1/2} \vee |\hat{D}(k)|. \quad (3.15)$$

Substitution into (3.12) yields the claim. □

### 3.2 Asymptotics for NBW on the torus

In this section, we study the convergence towards equilibrium of NBW on tori of width  $r \geq 3$ . We focus on two different examples. The first is random walk on products of complete graphs, where

$$\mathbb{V}_0 = \{x: \exists! i \in \{1, \dots, d\} \text{ such that } x_i \neq 0\}. \quad (3.16)$$

Our second example is NBW on the nearest-neighbor torus. The reason why we study these cases separately is that NBW on products of complete graphs is aperiodic, while nearest-neighbor NBW is periodic. Therefore, the stationary distribution for NBW on products of complete graphs equals the uniform distribution on the torus, while for nearest-neighbor NBW, the parity of the position after  $n$  steps always equals that of  $n$ . In Section 3.3, we further study random walk on the  $m$ -dimensional hypercube.

For any small  $\xi > 0$ , we write  $T_{\text{mix}}(\xi)$  for the  $\xi$ -uniform mixing time of NBW, that is,

$$T_{\text{mix}}(\xi) = \min \left\{ n: \max_{x,y} \frac{p_n(x,y) + p_{n+1}(x,y)}{2} \leq (1+\xi)V^{-1} \right\}, \quad (3.17)$$

where  $V = r^d$  is the volume of the torus. We start to investigate NBW on products of complete graphs:

**NBW on products of complete graphs.** Our main result is as follows:

**Lemma 3.3.** *For every  $d \geq 1$ ,  $r \geq 3$ ,  $n > \frac{d(r-1)}{r} \log((r-1)/[(1+\xi)^{1/d} - 1])$ , NBW on products of complete graphs satisfies that*

$$\max_{x \in \mathbb{T}_{r,d}} |p_n(x) - r^{-d}| \leq (m-1)^{-(n-1)/2} + \xi r^{-d}. \quad (3.18)$$

*In particular, for every  $\varepsilon > 0$ , there exists  $V_0$  such that when  $r^d \geq V_0$ ,*

$$T_{\text{mix}}(\xi) \leq (1+\varepsilon) \frac{d(r-1)}{r} \log((r-1)/[(1+\xi)^{1/d} - 1]).$$

When  $\xi$  is quite small, we obtain that  $T_{\text{mix}}(\xi) \leq (1 + 2\varepsilon) \frac{r}{r-1} \log(d(r-1)/\xi)$ .

*Proof.* The inverse Fourier transform on  $\mathbb{T}_{r,d}$  is given by

$$f(x) = \frac{1}{r^d} \sum_{k \in \mathbb{T}_{r,d}^*} \hat{f}(k) e^{ik \cdot x},$$

so that

$$b_n(x) = \frac{1}{r^d} \sum_{k \in \mathbb{T}_{r,d}^*} \hat{b}_n(k) e^{ik \cdot x} = r^{-d} \hat{b}_n(0) + \frac{1}{r^d} \sum_{k \in \mathbb{T}_{r,d}^* : k \neq \vec{0}} \hat{b}_n(k) e^{ik \cdot x}. \quad (3.19)$$

Therefore,

$$|p_n(x) - r^{-d}| \leq \frac{1}{r^d} \sum_{k \in \mathbb{T}_{r,d}^* : k \neq \vec{0}} \frac{|\hat{b}_n(k)|}{\hat{b}_n(0)}. \quad (3.20)$$

To bound  $|\hat{b}_n(k)|$ , we rely on Theorem 3.1, and start by computing

$$\hat{D}(k) = \frac{1}{d(r-1)} \sum_{i=1}^d \sum_{j=1}^{r-1} e^{ik_i j}. \quad (3.21)$$

Since  $k_i \in \frac{2\pi}{r} \{-\lfloor \frac{r-1}{2} \rfloor, \dots, \lceil \frac{r-1}{2} \rceil\}$ , we have that  $\sum_{j=0}^{r-1} e^{ik_i j} = 0$  for  $k_i \neq 0$ . Therefore,

$$\hat{D}(k) = \frac{1}{d(r-1)} \sum_{i=1}^d \sum_{j=1}^{r-1} e^{ik_i j} = \frac{1}{d(r-1)} \sum_{i=1}^d (r \mathbb{1}_{\{k_i=0\}} - 1) = 1 - \frac{r}{d(r-1)} a(k), \quad (3.22)$$

where  $a(k) = \sum_{i=1}^d \mathbb{1}_{\{k_i \neq 0\}}$  denotes the number of non-zero coordinates of  $k$ . By this observation,  $\hat{D}(k(j)) = 1 - \frac{rj}{d(r-1)}$  for any  $k(j) \in \mathbb{T}_{r,d}^*$  for which  $a(k(j)) = j$ . Then, by Theorem 3.1 and the fact that there are  $\binom{d}{j} (r-1)^j$  values of  $k \in \mathbb{T}_{r,d}^*$  for which  $a(k(j)) = j$ ,

$$\begin{aligned} |p_n(x) - r^{-d}| &= r^{-d} \sum_{j=1}^d (1/\sqrt{m-1} \vee |\hat{D}(k(j))|)^{n-1} \binom{d}{j} (r-1)^j \\ &\leq r^{-d} \sum_{j=1}^d \binom{d}{j} (r-1)^j \left[ \left| 1 - \frac{rj}{d(r-1)} \right|^{n-1} + (m-1)^{-(n-1)/2} \right] \\ &\leq (m-1)^{-(n-1)/2} + r^{-d} \sum_{j=1}^d \binom{d}{j} (r-1)^j e^{-rj(n-1)/[d(r-1)]}, \end{aligned} \quad (3.23)$$

where we use that  $|1 - \frac{rj}{d(r-1)}| \leq e^{-rj(n-1)/[d(r-1)]}$  for any  $j = 1, \dots, d$ . Thus,

$$\begin{aligned} |p_n(x) - r^{-d}| &\leq (m-1)^{-(n-1)/2} + r^{-d} [(1 + (r-1)e^{-r(n-1)/[d(r-1)]})^d - 1] \\ &\leq (m-1)^{-(n-1)/2} + \xi r^{-d}, \end{aligned} \quad (3.24)$$

when  $n > \frac{d(r-1)}{r} \log((r-1)/[(1+\xi)^{1/d} - 1])$ . The result on  $T_{\text{mix}}(\xi)$  follows immediately.  $\square$

**NBW on the nearest-neighbor torus.** Our main result is as follows:

**Lemma 3.4.** *For every  $d \geq 1$ ,  $r \geq 3$  and  $n > \log(2/[(1 + \xi/2)^{1/d} - 1])/[1 - \cos(2\pi/r)]$ , NBW on the  $d$ -dimensional nearest-neighbor torus satisfies that*

$$\max_{x \in \mathbb{T}_{r,d}} |p_n(x) - [1 + (-1)^{\|x\|_1 + n}]r^{-d}| \leq (2d - 1)^{-n/2} + \xi r^{-d}. \quad (3.25)$$

In particular, for every  $\varepsilon > 0$ , there exists  $V_0$  such that whenever  $r^d \geq V_0$ ,

$$T_{\text{mix}}(\xi) \leq (1 + \varepsilon) \log(2/[(1 + \xi/2)^{1/d} - 1])/[1 - \cos(2\pi/r)].$$

When  $r^d$  is large and  $\xi$  small, the above implies that  $T_{\text{mix}}(\xi) \leq (1 + 2\varepsilon)(r^2/(2\pi^2))(\log(2d/\xi))$ .

*Proof.* We adapt the proof of Lemma 3.3 to this setting. We have

$$b_n(x) = \frac{1}{r^d} \sum_{k \in \mathbb{T}_{r,d}^*} \hat{b}_n(k) e^{ik \cdot x} = r^{-d} [\hat{b}_n(0) + (-1)^{\|x\|_1} \hat{b}_n(\vec{\pi})] + \frac{1}{r^d} \sum_{k \in \mathbb{T}_{r,d}^* : k \neq \vec{0}, \vec{\pi}} \hat{b}_n(k) e^{ik \cdot x}. \quad (3.26)$$

Further,  $\hat{D}(-\vec{\pi}) = -1$ , so that, by (2.30)  $\hat{b}_n(\vec{\pi}) = \hat{b}_n(0)(-1)^n$ . Therefore,

$$\begin{aligned} |p_n(x) - [1 + (-1)^{\|x\|_1 + n}]r^{-d}| &= \left| \frac{1}{r^d} \sum_{k \in \mathbb{T}_{r,d}^* : k \neq \vec{0}, \vec{\pi}} \frac{\hat{b}_n(k)}{\hat{b}_n(0)} e^{ik \cdot x} \right| \leq \frac{1}{r^d} \sum_{k \in \mathbb{T}_{r,d}^* : k \neq \vec{0}, \vec{\pi}} \frac{|\hat{b}_n(k)|}{|\hat{b}_n(0)|} \\ &\leq (m - 1)^{(n-1)/2} + \frac{1}{r^d} \sum_{k \in \mathbb{T}_{r,d}^* : k \neq \vec{0}, \vec{\pi}} |\hat{D}(k)|^{n-1}. \end{aligned}$$

In the nearest-neighbor case,  $\hat{D}(k - \vec{\pi}) = -\hat{D}(k)$ , so we may restrict to  $k$  for which  $\hat{D}(k) \geq 0$ . Let

$$\mathbb{T}_{r,d,+}^* = \{k \in \mathbb{T}_{r,d}^* : \hat{D}(k) \geq 0\} \quad (3.27)$$

denote the set of  $k$ 's for which  $\hat{D}(k) \geq 0$ . Then,

$$|p_n(x) - [1 + (-1)^{\|x\|_1 + n}]r^{-d}| \leq (m - 1)^{(n-1)/2} + \frac{2}{r^d} \sum_{k \in \mathbb{T}_{r,d,+}^* : k \neq \vec{0}} \hat{D}(k)^{n-1}. \quad (3.28)$$

We use that

$$\hat{D}(k) = 1 - [1 - \hat{D}(k)] \leq e^{-[1 - \hat{D}(k)]}, \quad (3.29)$$

so that

$$\begin{aligned} |p_n(x) - [1 + (-1)^{\|x\|_1 + n}]r^{-d}| &\leq (m - 1)^{(n-1)/2} + \frac{2}{r^d} \sum_{k \in \mathbb{T}_{r,d}^* : k \neq \vec{0}} e^{-(n-1)[1 - \hat{D}(k)]} \\ &= (m - 1)^{(n-1)/2} + \frac{2}{r^d} \left[ \sum_{k \in \mathbb{T}_{r,d}^*} e^{-(n-1)[1 - \hat{D}(k)]} - 1 \right] \\ &= (m - 1)^{(n-1)/2} + \frac{2}{r^d} \left[ \left( \sum_{k \in \mathbb{T}_{r,1}^*} e^{-(n-1)[1 - \cos(k)]/d} \right)^d - 1 \right]. \end{aligned} \quad (3.30)$$

We use that the dominant contributions to the sum over  $k \in \mathbb{T}_{r,1}^*$  comes from  $k = 0$  and  $k = \pm 2\pi/r$ , so that

$$\sum_{k \in \mathbb{T}_{r,1}^*} e^{-(n-1)[1 - \cos(k)]/d} = 1 + 2e^{-(n-1)[1 - \cos(2\pi/r)]/d} (1 + o(1)), \quad (3.31)$$



so that

$$\begin{aligned} |p_n(x) - [1 + (-1)^{\|x\|_1 + n}]r^{-d}| &\leq (m-1)^{(n-1)/2} + \frac{2}{r^d} \left[ \left( 1 + 2e^{-(n-1)[1 - \cos(2\pi/r)]/d} (1 + o(1)) \right)^d - 1 \right] \\ &\leq (2d-1)^{-(n-1)/2} + \xi r^{-d} \end{aligned} \quad (3.32)$$

when  $n > (\log(2/[(1 + \xi/2)^{1/d} - 1])/[1 - \cos(2\pi/r)])$ .  $\square$

### 3.3 NBW on the hypercube

In this section we specialize the results of Sections 2 and Sections 3.1–3.2 to the hypercube  $\mathbb{T}_{2,m} = \{0, 1\}^m$ . The results in this section are an important ingredient to the analysis of percolation on the hypercube in [6]. We start with some notation. It will be convenient to let the Fourier dual space of  $\{0, 1\}^m = \{0, 1\}^m$  be  $\mathbb{Q}_m^* = \{0, 1\}^m$ , so that the Fourier transform of a summable function  $f: \{0, 1\}^m \rightarrow \mathbb{C}$  is given by

$$\hat{f}(k) = \sum_{x \in \{0, 1\}^m} f(x) e^{i\pi k \cdot x} = \sum_{x \in \{0, 1\}^m} f(x) (-1)^{k \cdot x}, \quad (k \in \{0, 1\}^m).$$

For  $k \in \{0, 1\}^m$ , let  $a(k)$  be its number of non-zero entries. Then, the SRW step distribution on  $\{0, 1\}^m$  satisfies

$$\hat{D}(k) = \frac{1}{m} \sum_{j=1}^m (-1)^{k_j} = 1 - 2a(k)/m. \quad (3.33)$$

The main result of this section is as follows:

**Theorem 3.5** (NBW on hypercube). *For NBW on the hypercube  $\{0, 1\}^m$ ,*

$$\hat{p}_n(k) \leq \left( |\hat{D}(k)| \vee 1/\sqrt{m-1} \right)^{n-1}. \quad (3.34)$$

Consequently, for every  $\varepsilon > 0$ , there exists  $m_0$  such that for all  $m \geq m_0$ ,

$$T_{\text{mix}}(\xi) \leq \frac{m(1 + \varepsilon)}{2} \log(2m/\xi). \quad (3.35)$$

Random walks on hypercubes have attracted considerable attention. In fact, the bound on the uniform mixing time for NBW in discrete time closely matches the one for SRW in continuous time (see [1, Lemma 2.5(a)]).

*Proof.* The bound in (3.34) follows directly from Theorem 3.1. We continue to investigate the convergence of the NBW transition probabilities  $x \mapsto p_n(x)$  to its quasi-stationary distribution, which is  $2 \times 2^{-m}$  when  $n$  and  $x$  have the same parity and 0 otherwise. Here we say that  $n$  and  $x$  have the same parity when there exists an  $n$ -step path from 0 to  $x$ .

**Lemma 3.6** (Convergence to equilibrium for NBW on hypercube). *For NBW on  $\{0, 1\}^m$  and every  $n > d(\log d + \log \xi)/2$ ,*

$$\max_{x \in \{0, 1\}^m} |p_n(x) - 2^{-m}[1 + (-1)^{\|x\|_1 + n}]| \leq (m-1)^{-(n-1)/2} + \xi 2^{-m}. \quad (3.36)$$

Consequently, for every  $\varepsilon > 0$ , there exists  $m_0$  such that for every  $m \geq m_0$ ,

$$T_{\text{mix}}(\xi) \leq -\frac{m(1 + \varepsilon)}{2} \log([1 + \xi/2]^{1/m} - 1).$$

*Proof.* The inverse Fourier transform on the  $m$ -dimensional hypercube is given by

$$f(x) = 2^{-m} \sum_{k \in \{0,1\}^m} \hat{f}(k) (-1)^{k \cdot x},$$

so that, using  $\hat{b}_n(\vec{1}) = (-1)^n \hat{b}_n(0)$ ,

$$b_n(x) = 2^{-m} \sum_{k \in \{0,1\}^m} \hat{b}_n(k) (-1)^{k \cdot x} = 2^{-m} [1 + (-1)^{\|x\|_1 + n}] \hat{b}_n(0) + \sum_{k \in \{0,1\}^m : k \neq \vec{1}, \vec{0}} \hat{b}_n(k) (-1)^{k \cdot x}. \quad (3.37)$$

Substituting the bound (3.34) in (3.37) leads to

$$\begin{aligned} |p_n(x) - 2^{-m} [1 + (-1)^{\|x\|_1 + n}]| &\leq 2 \cdot 2^{-m} \sum_{j=1}^{m/2} \binom{m}{j} [(1 - 2j/m)^{n-1} + (m-1)^{-n/2}] \\ &\leq (m-1)^{-n/2} + 2 \cdot 2^{-m} \sum_{j=1}^{m/2} \binom{m}{j} e^{-2j(n-1)/m} \\ &\leq (m-1)^{-n/2} + 2 \cdot 2^{-m} [(1 + e^{-2(n-1)/m})^m - 1] \\ &\leq (m-1)^{-n/2} + \xi 2^{-m}, \end{aligned} \quad (3.38)$$

when  $n > -\frac{m}{2} \log([1 + \xi/2]^{1/m} - 1)$ . □

**Acknowledgements.** This work was supported in part by the Netherlands Organization for Scientific Research (NWO). RvdH thanks Asaf Nachmias for valuable comments on an early version of the paper, and Gordon Slade for pointing us at the relevant literature.

## References

- [1] D. Aldous. Minimization algorithms and random walk on the  $d$ -cube. *Ann. Probab.*, **11**(2):403–413, (1983).
- [2] D. Aldous and J. Fill. *Reversible Markov Chains and Random Walks on Graphs* (monograph in preparation.), (2002).
- [3] N. Alon, I. Benjamini, E. Lubetzky, and S. Sodin. Non-backtracking random walks mix faster. *Commun. Contemp. Math.*, **9**(4):585–603, (2007).
- [4] N. Alon and E. Lubetzky. Poisson approximation for non-backtracking random walks. *Israel J. Math.*, **174**:227–252, (2009).
- [5] P. Billingsley. *Probability and Measure*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, third edition, (1995). A Wiley-Interscience Publication.
- [6] R. van der Hofstad and A. Nachmias. Hypercube percolation. Preprint (2012).
- [7] M. Kotani and T. Sunada. Zeta functions of finite graphs. *J. Math. Sci. Univ. Tokyo*, **7**(1):7–25, (2000).
- [8] D. Levin, Y. Peres, and E. Wilmer. *Markov Chains and Mixing Times*. American Mathematical Society, Providence, RI, (2009). With a chapter by James G. Propp and David B. Wilson.
- [9] N. Madras and G. Slade. *The Self-Avoiding Walk*. Birkhäuser, Boston, (1993).
- [10] J. Noonan. New upper bounds for the connective constants of self-avoiding walks. *J. Statist. Phys.*, **91**(5–6):871–888, (1998).
- [11] R. Ortner and W. Woess. Non-backtracking random walks and cogrowth of graphs. *Canad. J. Math.*, **59**(4):828–844, (2007).

- [12] A. Pönitz and P. Tittmann. Improved upper bounds for self-avoiding walks in  $\mathbf{Z}^d$ . *Electron. J. Combin.*, **7**:Research Paper 21, 10 pp. (electronic), (2000).
- [13] G. Slade. *The lace expansion and its applications*, volume **1879** of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, (2006). Lectures from the 34th Summer School on Probability Theory held in Saint-Flour, July 6–24, 2004, Edited and with a foreword by Jean Picard.